EXHIBIT F

From:	Steven Vachani
To:	Steven Vachani
Subject:	discussion with greg
Date:	Thursday, August 11, 2005 11:23:39 PM

Steve says:
good evening
greg says:
good evening
Steve says:
whats new
greg says:
preparing to launch
Steve says:
did I tell you that we rebuilt the orkut site in .net,
built an engine to grab the entire content of orkut
and store it in xml, and then let people's entire
content be activated instantly when they arrive on the
site
Steve says:
our nex step is to set up the servers appropriately to
deal witth traffic and numbers were discussing
Steve says:
also, when activating the engine to grab all the data,
we need to have a strategy that lets us grab all the
data without being detected
greg says:
 that sounds like it was your plan.
Steve says:
yeah... after we talked last, we were analyzing some
sites and made decision to switch to .net
Steve says:
and took also implemented your idea regarding
scalability
greg says:
well i never saw what he did in the end so i don't
know if it was really what i was describing or not
Steve says:
abi will probably send you a copy of the code shortly.
he redid the entire site in .net
Steve says:
do you have any specific ideas that would make it
easier to implement this engine we created to pull all
the content from orkut.
Steve says:
what I mean is the actual grabbing, not the engine
greg says:
that's sort of a wide open question. I don't really
know how the grabbing is happening even
Steve says:
we are pulling all the content for all 8 million
users, their scraps, photos, and most importantly,
their friends. stored in xml format in our database
and then the content will be activated when a user
enters our site
greg says:
grabbing data is a very technical thing, the problems
you'll run into will be very specific to how orkut is
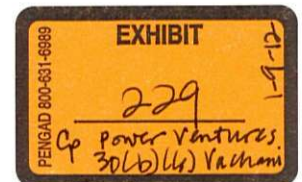
designed and what you try to do
greg says:
Pulling 8M users all at once will undoubtedly show up
as an increase in their bandwidth use too
Steve says:
yeah. we estimated that their is about 500k of data
per user on the system. we will have to have a clever
ip rotating system and other techniques. this is
something we are just starting to think about. they
have about 1 million users on the site per day
actively using it. so it will be interesting to figure
out how much bandwidth will go up and if the pattern
will be highly unusual
Steve says:
from alexa rankings, it looks like orkut has 1 billion
page views per day if I am reading correctly. so do
you think that it would be highly noticeabe if planned
correctly
greg says:
well if you download it from your servers the pattern
will be very unusual
greg says:
if you can somehow arrange for the user's browser to
be doing the download then it would be easier to miss
Steve says:
many users?
Steve says:
what do you mean somehow from a users browser
greg says:
if you get the user's browser to download the info
and then give it to you then at least the activity
would be spread amongst all the user's ip addresses
Steve says:
the challenge is how to get all the info initially
over a 1 week period. about 8 million users x 500k
greg says:
yeah i fear that may be doomed
Steve says:
and then only grab specific data when the user arrives
on our site and gives us their user name and
password... this allows us to activate their complete
profile in real time from our database
Steve says:
is their a way to user 25,000 different computers to
download all the data in some distributed project
greg says:
spammers do this, they distribute worms that infect
people's windows machines to do their work for them.
greg says:
I can't help you with that though
Steve says:
but aren't their some legal companies that pay to user
peoples computers
Steve says:
and they pay for use of people's computers
Steve says:
or their bandwidth
greg says:
not to do network things. especially not to do
distributed network attacks on one system

greg says:
akamai is kind of like what you describe, but i doubt
they would do something like this either. to do
something custom like this wouldn't be worth it for
them unless it was really lucrative.
Steve says:
there have projects in the past that have gone into
orkuts data to do analysis and research and
geolocational data, and other types of research. I am
not sure how much data they grabbed
Steve says:
assuming the ip address was dynamically changing, is
the danger the pattern that would be easy recognized?
Aren't their spiders from search engines that enter
these sites every day
Steve says:
and grab data
greg says:
If they notice the increase in bandwidth then they'll
look and see that an unreasonable percentage of their
overall usage is coming from your network. If they
don't notice then of course they don't notice. It's
not like orkut has a huge team on it, it's a side
project within google. They may not notice
greg says:
But if they notice it'll be pretty obvious what's
going on if it's all coming from your servers.
greg says:
That's why I'm saying the only way to really hide
things is to make the browser do the work for you. But
that could be hard
Steve says:
there are data mining/spidering companies out there
that specialize in this I think, right
greg says:
They don't download half a terabyte of data in a
single week
Steve says:
what kind of people or companies would have more info
on distributing and downloading a half a terabyte of
data through 5000 distributed computers. are there any
companies that would do a custom project like that
greg says:
I don't think any reputable company would do a project
like that
Steve says:
and who have access to 5000 computers who have
volunteered or made their comptuers available on a
paid basis
greg says:
You would want to speak to spammers to see what they
suggest
greg says:
You don't need 5000 computers, just 5000 ip addresses
greg says:
But that wouldn't really help you that much. orkut
would still be able to see that it's all the same
network unless it's something like i described with
the spammer worms or abusing open relays
Steve says:

hmmm.... sounds like a challenge. what about search
engine spiders which are grabbing this kind of data
spidering through orkut

greg says:

you keep asking the same questions, the answers aren't
going to change

greg says:

Those spiders go to a lot of effort *not* to download
this much data all at once

Steve says:

ok no problem. thanks. if you have any other ideas on
that, let me know

greg says:

precisely because if they don't they get noticed and
blocked

Steve says:

on another note. you thought about coming to brazil
any time soon? Second, is your schedule opening up any
time soon?

greg says:

the only solution i see short of using those kinds of
techniques spammers use is to use the users' browsers
to do the downloading. but then you can't do it until
the user shows up

greg says:

well we're preparing the launch now. I want to stick
around long enough to recoup my investment but I
should have more flexibility

Steve says:

i think that in the very near future, i would like to
find out how much time you will have available. if you
might have 10 hours a week of time available

Steve says:

or more?

greg says:

I should be able to do that in the near future

greg says:

not much more until we have more programmers on board
though

Steve says:

i have a new project, www.cupidgame.com launching this
month. it is a socal networking site where you invite
your friend to choose the best dates for you. we just
finished the functionality and are now doing a new
design before launch